Cell Stem Cell
Resource

# Efficient Endoderm Induction from Human Pluripotent Stem Cells by Logically Directing Signals Controlling Lineage Bifurcations

Kyle M. Loh,[1,2,11,*] Lay Teng Ang,[1,11,*] Jingyao Zhang,[1,12] Vibhor Kumar,[1,12] Jasmin Ang,[1] Jun Qiang Auyeong,[1] Kian Leong Lee,[3] Siew Hua Choo,[1] Christina Y.Y. Lim,[1] Massimo Nichane,[1] Junru Tan,[1] Monireh Soroush Noghabi,[1] Lisa Azzola,[4] Elizabeth S. Ng,[4,5] Jens Durruthy-Durruthy,[2] Vittorio Sebastiano,[2] Lorenz Poellinger,[3] Andrew G. Elefanty,[4,5] Edouard G. Stanley,[4,5] Qingfeng Chen,[6,7,8] Shyam Prabhakar,[1] Irving L. Weissman,[2] and Bing Lim[1,9,10,*]

[1]Stem Cell and Developmental Biology Group, Genome Institute of Singapore, Agency for Science, Technology and Research (A*STAR), Singapore 138672, Singapore

[2]Department of Developmental Biology, Stanford Institute for Stem Cell Biology and Regenerative Medicine, Stanford University School of Medicine, Stanford, CA 94305, USA

[3]Cancer Science Institute of Singapore, Centre for Translational Medicine, National University of Singapore, Singapore 117599, Singapore

[4]Department of Anatomy and Developmental Biology, Monash University, Clayton, VIC 3800, Australia

[5]Murdoch Childrens Research Institute, The Royal Children's Hospital, Parkville, VIC 3052, Australia

[6]Humanized Mouse Unit, Institute of Molecular and Cell Biology, A*STAR, Singapore 138673, Singapore

[7]Interdisciplinary Research Group in Infectious Diseases, Singapore-MIT Alliance for Research and Technology, Singapore 138602, Singapore

[8]Department of Microbiology, Yong Yoo Lin School of Medicine, National University of Singapore, Singapore 119228, Singapore

[9]Department of Medicine, Harvard Medical School, Boston, MA 02115, USA

[10]Division of Hematology/Oncology, Beth Israel Deaconess Medical Center, Boston, MA 02115, USA

[11]These authors contributed equally to this work

[12]These authors contributed equally to this work

*Correspondence: kyleloh@stanford.edu (K.M.L.), anglt1@gis.a-star.edu.sg (L.T.A.), limb1@gis.a-star.edu.sg (B.L.)

http://dx.doi.org/10.1016/j.stem.2013.12.007

## SUMMARY

**Human pluripotent stem cell (hPSC) differentiation typically yields heterogeneous populations. Knowledge of signals controlling embryonic lineage bifurcations could efficiently yield desired cell types through exclusion of alternate fates. Therefore, we revisited signals driving induction and anterior-posterior patterning of definitive endoderm to generate a coherent roadmap for endoderm differentiation. With striking temporal dynamics, BMP and Wnt initially specified anterior primitive streak (progenitor to endoderm), yet, 24 hr later, suppressed endoderm and induced mesoderm. At lineage bifurcations, cross-repressive signals separated mutually exclusive fates; TGF-β and BMP/MAPK respectively induced pancreas versus liver from endoderm by suppressing the alternate lineage. We systematically blockaded alternate fates throughout multiple consecutive bifurcations, thereby efficiently differentiating multiple hPSC lines exclusively into endoderm and its derivatives. Comprehensive transcriptional and chromatin mapping of highly pure endodermal populations revealed that endodermal enhancers existed in a surprising diversity of "pre-enhancer" states before activation, reflecting the establishment of a permissive chromatin landscape as a prelude to differentiation.**

## INTRODUCTION

At developmental junctures, multipotent progenitors choose between multiple fates (Graf and Enver, 2009; Loh and Lim, 2011). Extrinsic signals often instruct a particular fate while repressing alternate lineages. It is critical to decipher the extrinsic signals that direct such lineage segregations in order to efficiently differentiate human pluripotent stem cells (hPSCs) into pure populations of desired cell types in the absence of mutually exclusive, unwanted lineages. However, the precise lineage outcomes specified by these signals at particular bifurcations remain to be fully clarified, despite informative insights from in vivo genetic perturbations (Tam and Loebel, 2007; Zorn and Wells, 2009) and explant approaches (Bernardo et al., 2011; Deutsch et al., 2001). Pertinent issues include how alternate lineages are segregated at each branchpoint as well as the exact order and kinetics of dynamic signaling switches that drive successive cell fate transitions (Wandzioch and Zaret, 2009).

The present work revisits signaling dynamics that drive induction and anterior-posterior patterning of the definitive endoderm (DE) germ layer and subsequent organ formation. DE is the embryonic precursor to organs including the thyroid, lungs, pancreas, liver, and intestines (Švajger and Levak-Svajger, 1974). The pluripotent epiblast (E5.5 in mouse embryogenesis) differentiates into the anterior primitive streak (E6.5), which generates DE (E7.0–E7.5) (Lawson et al., 1991; Tam and Beddington, 1987). DE is then patterned along the anterior-posterior axis into distinct foregut, midgut, and hindgut territories (E8.5), and endoderm organ primordia arise

from specific anteroposterior domains (E9.5) (Zorn and Wells, 2009).

Various methods to differentiate hPSCs toward DE employ animal serum, feeder coculture, or defined conditions (Cheng et al., 2012; D'Amour et al., 2005; Touboul et al., 2010), but they typically yield a mixture of DE and other contaminating lineages, with induction efficiencies fluctuating between hPSC lines (Cohen and Melton, 2011; McKnight et al., 2010). Viewed from the perspective of lineage bifurcations, these mixed lineage outcomes might stem from incomplete exclusion of alternate fates at such junctures. Heterogeneous early DE populations harboring contaminating lineages complicate the subsequent generation of endodermal organ derivatives (McKnight et al., 2010).

In vertebrate embryos and during PSC differentiation, TGF-β/nodal/activin signaling is imperative for DE specification, whereas BMP broadly induces mesodermal subtypes (e.g., Bernardo et al., 2011; D'Amour et al., 2005; Dunn et al., 2004). Yet, TGF-β signaling (even with additional factors) is insufficient to specify homogeneous DE (quantified by Chetty et al., 2013). BMP, fibroblast growth factor (FGF), VEGF, and Wnt have also been employed together with TGF-β signals to generate DE (Cheng et al., 2012; Green et al., 2011; Kroon et al., 2008; Nostro et al., 2011; Touboul et al., 2010). However, these factors have also been implicated in mesoderm formation (Davis et al., 2008), and their precise involvement in DE induction remains to be clarified.

We have systematically elucidated how mutually exclusive lineages are separated at four consecutive steps of endoderm development: PS induction, segregation of endoderm versus mesoderm germ layers, DE anterior-posterior patterning, and bifurcation of liver and pancreas. Accurately defining which signals instructed or repressed specific fates at each endodermal bifurcation enabled homogeneous hPSC differentiation down one path or the other. Knowledge of precise temporal signaling dynamics, combined with efficient differentiation throughout successive developmental steps, culminated in a single strategy to universally differentiate diverse hPSC lines into pure populations of endodermal lineages by excluding alternate lineages at each branchpoint. Altogether, this provides a coherent view of signaling logic underlying multiple steps of endoderm induction and patterning. This also furnishes the means to molecularly profile highly homogeneous endoderm populations, allowing us to comprehensively capture transcriptional and chromatin dynamics underlying endoderm specification.

## RESULTS

### BMP, FGF, TGF-β, and Wnt Initially Establish the Primitive Streak and Anteroposteriorly Pattern It

This work was preceded by findings that activin, in conjunction with FGF, BMP, and a phosphatidinositol 3-kinase (PI3K) inhibitor ("AFBLy") (Touboul et al., 2010) or together with animal serum (D'Amour et al., 2005), induced hESCs toward DE. However, we and others (Chetty et al., 2013) observed that these methods still yielded mixed lineage outcomes, which was evident during the differentiation of five hESC lines (Figures 1A and 2B and 2C; Figures S1–S3 available online). For example, AFBLy (Touboul et al., 2010) concurrently generated mesoderm, upregulating skeletal, vascular, and cardiac genes ($p < 10^{-8}$; Figure 1A; Figures S1A–S1D), whereas activin and serum treatment (D'Amour et al., 2005) yielded a proportion of undifferentiated cells (Figures 2B, 2C, and 2F). Creation of impure early DE populations might explain the emergence of nonendoderm lineages after downstream differentiation (Kroon et al., 2008; Rezania et al., 2012).

Guided by prior in vivo and in vitro findings, we selectively perturbed developmental signals (>3,200 signaling conditions) at specific embryonic stages of hPSC differentiation in serum-free conditions and assessed resultant lineage outcomes by qPCR (yielding >16,000 data points, Figure S1–S4). These signaling perturbations revealed elements of the signaling logic underlying DE induction (Figures 1, 2, 3, and 4).

In vivo, DE arises from the primitive streak (PS, E6.5) (Lewak-Svajger and Švajger, 1974). The anteriormost PS (APS) generates DE (E7.0–E7.5), whereas posterior PS (PPS) forms mesoderm (Lawson et al., 1991; Tam and Beddington, 1987). Determinants of anterior versus posterior PS from hPSCs remain to be elucidated.

We found both APS and PPS were combinatorially induced by BMP, FGF, and Wnt on day 1 of hESC differentiation. These signals have been individually implicated in PS induction (Bernardo et al., 2011; Blauwkamp et al., 2012; Gadue et al., 2006), but their roles in PS patterning have not been dissected in detail. If BMP, FGF, or Wnt were inhibited, both APS and PPS formation failed (Figure 1B), corroborating the lack of PS in BMP and Wnt pathway knockout mice (Beppu et al., 2000; Liu et al., 1999; Mishina et al., 1995). FGF signaling was equally permissive for both APS and PPS emergence, and endogenous FGF was sufficient to drive either outcome (Figure 1Bi, Figures S2A–S2C). However, exogenous Wnt (either Wnt3a or GSK3 inhibition [CHIR]) was necessary to maximize PS induction, and Wnt broadly promoted both APS and PPS (Figure 1Bii and 1Biii). Limited PS formation could occur without exogenous Wnt, but was dependent on endogenous Wnt (Figure 1Bii). BMP levels arbitrated between APS and PPS; lower (endogenous) BMP levels elicited APS, whereas higher BMP yielded PPS (Figure 1Biv; Figure S2B). Nonetheless, the absolute necessity of BMP for *MIXL1-GFP*$^{+}$ APS induction (Figure 1Di, $p < 0.025$) was unexpected, because BMP was typically thought to be posteriorizing (Bernardo et al., 2011). Therefore, FGF, Wnt, and low BMP were essential for APS specification.

### A Dynamic Switch in BMP and Wnt Signaling Induces Primitive Streak but Subsequently Suppresses DE Emergence

To further differentiate APS toward DE, prior studies used similar factors to induce both lineages over 3–5 days (Nostro et al., 2011; Touboul et al., 2010). Instead, we found that APS and DE were sequentially driven by diametrically opposite signals within 24 hr of differentiation. BMP and Wnt initially specified APS from hESCs on day 1, but, 24 hr later, BMP and Wnt induced mesoderm and reciprocally repressed DE formation from PS on days 2–3 (Figures 1Ci and 1Cii). Interestingly, not only removing exogenous BMP but neutralizing endogenous BMP (using noggin or DM3189/LDN-193189) was critical to eliminate mesoderm and to reciprocally divert PS differentiation toward DE (Figure 1Ci). This was evinced by 3,000-fold downregulation of *MESP1* and concurrent upregulation of *SOX17*, *HHEX*, *FOXA1*, and *FOXA2* in two hESC lines (Figures S1C–S1E). Given that